MRC# 78

Free Voice and Pure Tone Audiometer for Routine Testing of Auditory Acuity

Studies on Comparative Efficiency

J. DONALD HARRIS, Ph.D.
NEW LONDON, CONN.

Reprinted from the Archives of Otolaryngology
October 1946, Vol. 44, pp. 452-467

Printed and Published in the United States of America

Reprinted from the Archives of Otolaryngology
October 1946, Vol. 44, pp. 452-467

FREE VOICE AND PURE TONE AUDIOMETER FOR ROUTINE TESTING OF AUDITORY ACUITY

Studies on Comparative Efficiency

J. DONALD HARRIS, Ph.D.
NEW LONDON, CONN.

THE RELATION between a free voice acuity test and a pure tone audiogram is a complex one of many factors; no final statement has yet been made by which the interpretation of either can be made in terms of the other or by which both can be reduced to a common denominator. Some laboratories depend exclusively on a pure tone audiogram as a measure of how well a man can hear speech, while others look askance on the pure tone test and claim that only a test using speech sounds is valid as a measure of speech reception.

An early discussion of the relation between speech reception and audiometry is that of Fletcher,[1] who compared a free voice test and audiometry through the speech range (512 to 2048 cycles per second). Fletcher concluded, from the rather close correspondence of hearing loss, that the free voice and the pure tone audiogram measure essentially the same auditory function. It is clear that on the basis of laboratory data there is empiric justification for attempts to equate the two tests in routine situations.

However, when one comes to examine the relationship between the two measures as routinely obtained, one finds that in some cases little correspondence may exist. When, for example, in one experiment [2] the voice was placed at the low conversation intensity level, the authors concluded, "On the basis of our results, the spoken voice test may be summarily dismissed as of no diagnostic value."

But even when the voice is placed at the whispered intensity level, the screening out of persons with poor acuity through the speech range may be far from satisfactory, and, under conditions which generally prevail as to test chamber and training of personnel, the value of the whispered voice test may be expected to fall off sharply from maximum.

In any situation in which it is necessary to have the best available measure of acuity for speech, the worker, in deciding what test to use,

---

From the Medical Research Laboratories, Submarine Base.

1. Fletcher, H.: Speech and Hearing, New York, D. Van Nostrand Company, Inc., 1929, chap. 6.

2. Shilling, C. W.; Everley, I. A., and Harris, J. D.: Hearing Tests: An Evaluation, U. S, Nav. M. Bull. **44**:100-116, 1945.

must consider such factors as the sort of sound-treated space available, the amount of testing time per man and the reliability and significance of the tests which it is possible to administer in the available space and time.

The present reference paper is designed to help the beginning worker survey his equipment and his needs and reach a decision as to which test or tests he should institute and to provide comparable data against which an experienced worker may check his own procedures.

The practical questions arise: (1) whether the free voice and pure tone audiometry actually are equivalent as measures of threshold acuity; (2) whether in a particular situation, with physical conditions less carefully controlled than in a laboratory, one test can be administered with more reliability and efficiency and so is to be preferred, and (3) if the two tests are of equal reliability, which test should be used to fit best the purposes of a particular activity. These questions will be considered in that order.

I. DO THE FREE VOICE AND PURE TONE AUDIOMETRY THROUGH THE SPEECH RANGE MEASURE THE SAME THING?

An extensively documented paper by Goldman[3] recently stated, "An analysis of the audiometric and whisper readings revealed that there was no constant correlation between the audiometric average decibel loss and the recording of the whisper perception. The same deficiency was noted between the audiometric and the low conversational voice tests." In view of such broad statements it is necessary at this time to examine the evidence on which Fletcher earlier concluded that the two tests were related and to examine later evidence bearing on the same point.

Fletcher's tables XXIV and XXV, pages 219 and 220, provide a comparison of hearing loss in decibels for 18 ears, first according to the free voice and again according to the average audiometric loss through the speech range. My colleagues and I have computed the correlation between hearing loss according to Fletcher's tests to be 0.95 ± 0.02. This is certainly a high degree of relationship. The standard error of estimate is 5.4 decibels, which means that by knowing the loss on either test one can predict loss on the other within 5.4 decibles about two thirds of the time. These figures provide strong grounds for the suppostion that the free voice and audiometry in the same range have not only a demonstrable but a high relationship when both are administered with laboratory care.

The trend of Fletcher's data for 18 ears is confirmed by a study from this laboratory on 335 ears, of which 126 scored below normal on the

3. Goldman, J. L.: A Comparative Study of Whisper Tests and Audiograms, Laryngoscope **54**:559-572, 1944.

whispered voice test. Thus the data need not be corrected for a restricted range of ability. The correlation between the whispered voice and the average audiogram loss from 512 to 2048 cycles per second was 0.69. All tests were conducted under less well controlled physical surroundings than Fletcher's data and under routine rather than experimental conditions, and the resultant drop in reliability could easily account for the drop of 0.26 in the correlation. Nevertheless the correlation of 0.69 is considered fairly high for such material, and it is probable that it is about as high a relationship as routine conditions will permit.

A more reliable voice test is one in which phonograph recordings replace the tester's own voice. In this case, hearing loss is measured directly in decibels rather than calculated from the voice distance fraction. One can find the relation between hearing loss according to the Western Electric 4-A phonograph audiometer and according to pure tone audiometry from Fletcher's same two tables. The standard error of estimate is 4.7 decibels. Evidently the relationship of the audiogram to the phonograph score is 0.7 decibels closer than to the whispered voice score. This increase in precision is almost certainly due to the increased reliability of the phonograph over the whispered voice.

Ciocco [4] likewise compared the average 4-A phonograph score with the average audiometer score at 1024 and 2048 cycles per second. For all groups whose loss averages between 0 and 50 decibels on the audiogram, a plot relating the 4-A score to the audiometer score can adequately be described as a straight line. An exact correspondence between absolute scores is not the case; rather, the phonograph score is approximately 0.6 that of the audiogram, but it is the linearity of the relationship which is important for the present argument.

In this laboratory, we have compared a series of 180 ears on pure tone audiometry and on a new recorded test of speech reception.[5] Among these 180 ears are found samples at all levels of acuity. A comparison of hearing loss according to these records and the average audiometer loss for 512 to 2048 cycles per second confirms Fletcher's experience in all details. On 94 ears with audiograms flat $\pm$ 5 decibels, we have found that the correlation between phonograph and audiometer scores is 0.93, with a standard error of estimate of 3.9 decibels. But even with 83 ears in which the audiogram varied more than 10 decibels (presumably those ears in which amplitude distortion would reduce intelligibility) we found a standard error of estimate of only 5.8 decibels.

---

4. Ciocco, A.: The Consistency and Significance of Tests Made with a 4 A Audiometer, Pub. Health Rep. **51**:1402-1406, 1936.

5. Manual of Instruction for Auditory Test No. 9: Threshold of Hearing for Words, Informal Communication number 73, NDRC Research on Sound Control, Psycho-Acoustic Laboratory, Harvard University, May 20, 1944.

Instead of the phonograph, a monitoring system for the voice may be used to control its intensity. The operator talks into a microphone, keeping his voice at constant intensity by watching a voltmeter and decreasing the output to the subject's loudspeaker by inserting resistance into the circuit. This technic was reported by Hughson and Thompson.[6] When their data are replotted in absolute terms, a precise linearity appears between speech reception and audiometry. Not only do their data plot in straight line fashion, but there is the closest correspondence between absolute scores—a loss of 50.3 decibels for speech is related to an average audiometer loss of 48.2 decibels and so on.

It has been thus seen that whether one considers the free voice (Fletcher and this laboratory), the phonograph voice (Fletcher and Ciocco, this laboratory) or the monitored voice (Hughson and Thompson) in general it is true that reception for speech and for pure tones in the speech range maintain a close parallel.

There is a certain error of prediction, which has been shown to be of the order of a few decibels, and there are one or two clinical types for which the correspondence may not hold but these considerations can hardly account for the cases in which speech and pure tone losses do not fairly well agree. In those cases one may reasonably look for some factor making for unreliability of one or both the tests. It is almost certain that in cases in which a worker finds a correlation lower than about 0.70 his speech reception scores or his audiometry data are less reliable than those possible even in routine military conditions such as ours. For 101 nerve-deaf ears reported by Goldman,[3] for example, the correlation between free voice and average audiometer loss at 256 to 2048 cycles per second was 0.24. But it can easily be shown (see figure 1) that in the test alley used by Goldman the tester had to move from 15 feet (4.5 meters) to within 2½ feet (0.7 meter) of the subject's ear before a loss of 5 decibels could be detected, while the maximum loss that could be detected even from the 6 inch (15 cm.) distance was about 15 decibels. His speech test, therefore, was not capable of wide enough range. We may agree with Goldman that in his test alley there was little correlation between speech reception and audiometry, but it is not necessary to apply this conclusion to other test situations.

Macfarlan[7] likewise stated, "There is no correlation between speech hearing and frequency hearing." However, when his data are plotted, it turns out that a correlation of 0.65 is present, while if his audiogram

6. Hughson, W., and Thompson, E.: Correlation of Hearing Acuity for Speech with Discrete Frequency Audiograms, Arch. Otolaryng. **36**:526-540 (Oct.) 1942.

7. Macfarlan, D.: Speech Hearing and Speech Interpretation Testing, Arch. Otolaryng. **31**:517-528 (March) 1940.

scores are multplied by 0.6 (the same factor found for Ciocco, it will be recalled) in only 6 of the 21 ears reported will the prediction of one score from the other be erroneous by more than 10 decibels. These are not striking figures, but at any rate it does seem from his own data that Macfarlan's conclusion of "no correlation" is premature.

The general conclusion to be drawn from the foregoing data is that for typical subjects and in cases in which vocabulary is controlled speech reception and pure tone audiometry are intimately connected and are concerned with measuring practically the same auditory function.

II. UNDER CONDITIONS OF TESTING WHICH ACTUALLY PREVAIL, WHAT ARE THE LIMITS OF RELIABILITY AND EFFICIENCY OF WHICH THE FREE VOICE AND AUDIOMETRY ARE CAPABLE?

The conclusions of the first section of this paper are based largely on laboratory conditions. What must be discovered is how much reduction in reliability occurs when these tests are taken out of the laboratory and placed in the hands of routine testers under inferior and unsupervised test conditions.

A. The Reliability of Audiometry.—First the reliability of audiometry must be considered. Munson [8] compared thirty-eight test and retest scores on the Western Electric 6-A audiometer and reported the standard deviations of the differences between scores. He found that there were two chances in three that the retest would be within 4.3, 4.0 and 4.3 decibels for the respective frequencies 512, 1024 and 2048 cycles per second.

These figures, taken as they were, in a good laboratory under favorable conditions, may be taken as the precision of which a test using steps of 5 decibels intensity is capable. However, results as good as these have been obtained under service conditions. Giese,[9] on 109 subjects, reported standard deviations ranging from 3.6 to 4.6 decibels for the same frequencies as Munson. Harris,[10] for 64 subjects, reported data not differing essentially from these.

In terms of test-retest correlations, Giese's data show reliabilities from 0.63 to 0.73 and Harris' data from 0.70 to 0.87. In both of these groups these correlations are lower than would have been the case had a greater range of acuity been present. Even so, the data of Giese and of Harris compare favorably with the data of Witting and Hughson [11] on

8. Munson, W. A.: Trial Tests of Pulsing Tone Audiometer, unpublished research memorandum, Case 20871-2, New York, Bell Telephone Laboratories, Inc., 1937.

9. Giese, W. J.: Test-Retest Reliability of the Western Electric 6 B Audiometer Under Military Conditions, OSRD Report no. 91.11, Aug. 30, 1943.

10. Harris, J. D.: Group Audiometry, J. Acoust. Soc. America **17**:73-76, 1945.

11. Witting, E. G., and Hughson, W.: Inherent Accuracy of Repeated Clinical Audiograms, Laryngoscope **50**:259, 1940.

repeated audiograms. The latter's standard deviations of repeated audiograms on the same person (from 2.5 to 4.4 decibels) can be indirectly compared with data on only two audiograms but for more ears.

It may be concluded that pure tone audiometry as a general method, when performed by experienced operators under good room conditions with an audiometer in calibration and performing well, can easily satisfy minimum requirements as to reliability and as to uniformity of interpretation and that this may be accomplished in military situations.

When the retest is administered, in a second room, with a second tester using a different make of audiometer, the correspondence of test-retest will be somewhat less. We have made a comparison of audiograms taken routinely in our laboratory by group audiometry, using the Western Electric 6-B instrument, with audiograms on the same 192 ears taken routinely on a Maico audiometer about three months later at the West Coast Sound School, San Diego, Calif. The latter audiograms were furnished by Dr. Adelbert Ford. Although the mean acuity loss was about 10 decibels greater through the speech range for the data from San Diego (due in all probability to our more nearly sound-proof test chamber), nevertheless the correspondence between the two sets of data is satisfactory. In terms of the standard error of estimate, there were two chances in three of predicting the loss recorded at San Diego for any ear by 4.7, 4.3 and 5.3 decibels at 512, 1024 and 2048 cycles per second respectively.

The training and experience of the operator contribute somewhat to the reliability of an audiogram. For the most complete audiometry, involving bone conduction, localization and the like, only an operator with months of study under a good clinician will perform satisfactorily. But for routine audiometry, as, for example, in military screening, where most of the population is of normal or near-normal hearing, extensive training is not necessary. A study of test-retest reliability of three enlisted operators at this activity, where the retest was taken under the same conditions as the test (the operator being unaware that a retest was to be called for and a period of time elapsing to reduce incidental learning), shows that an operator can early attain a certain stability. In all cases these operators had had a week's informal lectures and demonstrations, had read most of Bunch's "Clinical Audiometry,"[12] part 3 of Fletcher's "Speech and Hearing"[1] and chapters 1, 2, 10, 11 and 12 of Stevens and Davis' "Hearing."[13] The table shows how much average deviation exists between test and retest scores for a

12. Bunch, C. C.: Clinical Audiometry, St. Louis, C. V. Mosby Company, 1943.

13. Stevens, S. S., and Davis, H.: Hearing, New York, John Wiley & Sons, Inc., 1938.

sophisticated operator and for each of three operators given only a minimum of training and experience.

The procedure used by the experienced operator was one in which only two crossings of the limen were used in determining the final score. This procedure was taught to each operator as a minimum requirement; in case of doubt he was instructed to repeat readings. It is clear that there is a decrease in precision of 1 or 2 decibels for the enlisted operators but that with only slight training practically any intelligent person can learn to do audiometry on routine near-normal subjects with almost as good accuracy as a trained tester. It is my experience that more confidence can be placed in an audiogram taken by a technician with known personality traits of patience, accuracy and ability to put subjects at ease than in an audiogram taken by one whose chief qualification is that he has had a large amount of routine practice.

The general conclusion to be reached as to the limits of reliability of which audiometry is capable is that where space is provided with

Mean Deviation of Retest from Test Score, Western Electric 6-B Audiometer Being Used

Cycles per Second	Experienced Operator	Operator 1	Operator 2	Operator 3
512	2.30	3.40	4.50	3.40
1,024	2.00	4.15	4.05	2.35
2,048	2.10	4.85	4.25	2.80

reasonably low levels of noise only briefly trained technicans can furnish audiograms accurate to at least 5 decibels.

B. The Meaning and Reliability of the Free Voice Test.—In the case of the free voice test, the reliability statistics are not quite so encouraging. It is notorious that in cases in which a doctor is anxious that his therapy have some effect he is liable to the natural error of unconsciously raising his intensity level on a voice retest. But especially with routine testers the situation is critically bad. As often administered in the services, for example, the voice test screens out only the men with bilateral loss severe enough to require a hearing aid for normal conversation.

Of course, the fact that a test may be grossly misused does not mean that the test itself is valueless, but it must be emphasized that the free voice acuity test contains at best many sources of unreliability and that in cases in which conditions are not good its value even for rough screening is problematic.

It will serve no purpose to dwell on instances in which, owing to lack of supervision, the free voice test failed to detect a deafened ear.

What one needs to know is whether, given reasonable supervision, routine testers can perform satisfactorily. A measure of just how

satisfactory such testers can be is provided by data from this activity, in which a careful whispered voice test is given to all candidates for the submarine service. Every man sent to the submarine base has received a whispered voice score of 15/15 at the immediately preceding activity. This previously given test was administered under conditions as good as could be secured by frequent visits and communications from the officer in charge [14] of this activity. In these circumstances, the subsequent whispered voice test given here failed 68 of 1,555 supposedly qualified candidates, or 4.3 per cent, in a five month period. This figure represents a rather close correspondence between test and retest at different stations. From data obtained earlier, we know that had the whispered voice test not been given at the preceding station or been given in poorer fashion many more than 68 candidates would have had to be rejected for low auditory acuity.

For example, in one station from which the rejections at this activity were 5.7 per cent for a four month period, the whispered voice testing was found to be carried on in a reverberative metal Quonset hut; when this condition was corrected, our rejection of candidates from that station fell off to less than 1 per cent in the succeeding three month interval.

A check on the whispered voice test as conducted at this activity is provided by a careful 6 octave audiogram given routinely to all candidates for submarine service. For a typical three month period, out of 5,248 ears our whispered voice test labeled as normal 92 ears which had an average loss of 15 or more decibels through the speech range. Since an average loss of 15 decibels should correspond to a loss of several feet on an ideal voice test, these 92 ears, or 1.75 per cent, are an index of the error of selection of our voice test.

These data on the efficacy of the whispered voice show it to be of definite value as a quick screening device but show that even under the best service conditions an appreciable number of defective ears are overlooked.

In an effort to make the whispered voice score more meaningful, an experiment was undertaken to determine the "normal" distance for two types of phonetic material and for three types of test rooms. Five pharmacist's mates were used as testers, and 13 enlisted men with normal audiograms $\pm$ 5 decibels were used as subjects. Tests were conducted in a soundproof room 18 by 13 by 9 feet (5.4 by 3.9 by 2.7 meters) high, with attenuation from the outside in excess of 90 decibels; a classroom 29 by 21 by 10 feet (8.8 by 6.4 by 3 meters), with an average level of 50 decibels measured with a sound level meter, and, finally, out-of-doors, with sound level estimated at 35 to 40 decibels.

14. Captain C. W. Shilling (MC), U.S.N., Officer-in-Charge, Medical Research Laboratories.

The following tests were administered to each experimental ear by each tester in each room; in most instances a retest was run under so far as possible the same conditions:

1. Whispered voice, only the words "one," "two," "three," "four," "five," "six" and "eight" being used. Each tester was given directions consisting of those in the Manual of the Medical Department, U. S. N., 1939, plus the directions that at any distance the subject must repeat two out of three pairs of words correctly (the words to be said in pairs, thus: "one-three," "four-five," etc.) before the subject is given credit for that distance.

2. Whispered voice, a group of eighty-four spondees which have been shown to be of equal intelligibility being used.[5]

All the usual requirements of good experimental design and procedure were followed; testers were not told that all ears were normal in acuity or that their averages would be obtained; record sheets were collected from each tester as soon as he had finished with any subject on any test and similar requirements were fulfilled.

The results show that the designation of 15 feet (4.5 meters) as the denominator in the whispered voice fraction is approximately correct for typical Navy testing. For the digit list the test and retest mean distances in the classroom were 14.6 and 15.6 feet (4.4 and 4.7 meters) respectively, in the sound-proof room the distances were 16.1 and 15.8 feet (4.9 and 4.8 meters) and out-of-doors the mean test distance was 14.3 feet (4.3 meters). In these figures can be seen the tester's tendency to place his voice level in adjustment to the noise surround.

The same tendency can be observed in the case of the spondee list, for which in the sound-proof room the test and retest means were 14.9 and 14.8 feet (4.54 by 4.51 meters), while in the noisy classroom the retest mean was 14.2 feet (4.3 meters).

The effects of vocabulary and of learning in the whispered voice test are demonstrated by a difference in mean distance between test and retest scores for the spondee list in the classroom. On first exposure to these words the mean score was 11.3 feet (3.4 meters) (4.8 feet [1.5 meters] shorter than the first test using digits), but by the second test this distance was increased to 14.2 feet (4.3 meters), a figure agreeing well with the other data.

It may be concluded that a distance of 15 feet (4.5 meters) will, for a variety of test conditions and as an average of several testers, be found close to the "normal" distance.

To determine how close a tester can approximate a previous score on the same ear, one takes the average deviation[*] of retest from test scores. The data from the classroom show that the chances are two in three that a tester can come within 3.7 feet (1.1 meters) for digit material and within 2.4 feet (0.7 meter) for the spondees when learning

is eliminated. The reliabilities for the sound-proof room are uniformly a little better.

A standard error of estimate of about 3 feet (0.9 meter) then expresses the precision of which this test is capable in routine hands. When expressed in decibels (see following material), it can be compared with the same datum for audiometry reported earlier in this paper.

Although every effort was made to simulate routine screening, the conditions of this experiment were nevertheless considerably better than ordinary ones, in the training of personnel, enthusiasm for testing and general supervision. Still, it is certainly fair to regard the precision reported here as within the reach of routine acuity testing.

The standard error of estimate of about 3 feet translates into only a few decibels (see figure 1) even for ideal room conditions. This compares favorably with the error reported here for audiometry and indicates that in cases in which supervision and test conditions approach ideal the acuity of normal or near-normal subjects can be placed with reasonable precision in routine screening.

C. Hearing Loss and the Acoustics of Sound Decay.—It is not possible any longer to defer a discussion of the error of predicting from the free field equation for hearing loss what will be the results from an actual testing chamber. According to the law that sound pressure falls off inversely as the square of the distance from the source, hearing loss $= 20\ (\log d_0 - \log d)$ where d_0 is the distance at which a normal ear can just hear the sound used and d is the distance to be considered. But this equation is derived for a perfectly reflectionless room, which most acuity testing chambers assuredly are not. What is needed is an idea of the way in which sound actually does decay as a function of distance in typical testing chambers.

We have obtained an estimate of the decay of sound for three representative types of test rooms. The index used was the number of decibels a voice intensity had to be raised to maintain constant intelligibility as the sound moved away from the ear.

The specific procedure was as follows: A loudspeaker was set up 15 feet (4.5 meters) from a normal-hearing ear, and discrete words were presented by means of phonograph recording. An attenuation pad in 1 decibel steps was inserted between the phonograph output and the loudspeaker. The percentages of intelligibility for several attenuation readings were determined, and then the psychophysical function was drawn relating per cent intelligibility to attenuation reading; finally, from this curve, the attenuation reading corresponding to 75 per cent intelligibility was determined.

At this point the loudspeaker was moved to within 13 feet (3.9 meters), and again the attenuation reading corresponding to 75 per cent

intelligibility was determined. This process was repeated (suitably counterbalanced to control learning) for distances up to 6 inches (15.2 cm.).

The difference in attenuation readings for 15 feet and any other distance is found in decibels. It can be thought of as a measure of decay of ordinary speech through that particular distance and so can be compared with the theoretic computation of sound decay in a free field. Any discrepancy between a datum according to our method and the theoretic calculation for the same distance may be explained in terms of the lack of free field conditions in the particular test room used.

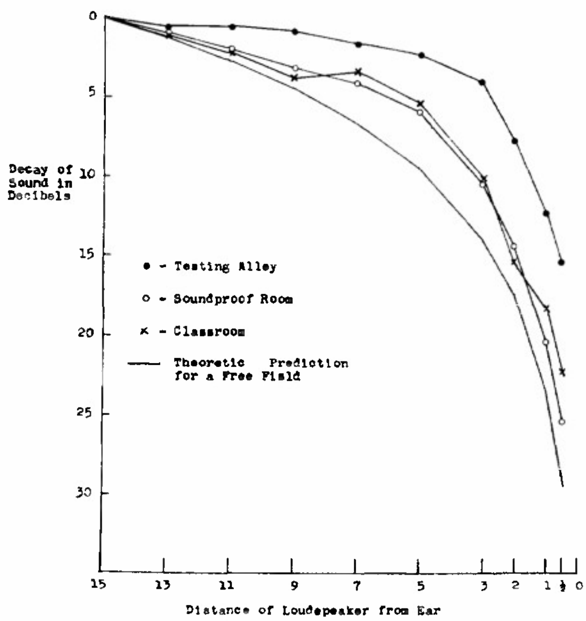

Fig. 1.—Sound decay as a function of distance.

It is of interest to note that the datum described here can also be taken as the decibel hearing loss corresponding to a whispered voice distance fraction. As the sound source moves from 1 to 15 feet (0.3 to 4.5 meters) from the ear, it must be made so many decibels louder to maintain constant intelligibility. But it is equally correct to say that an ear deafened by that same amount will have to be within 1 foot of a sound which a normal ear could hear at 15 feet. We have in this technic and in this chain of reasoning a rational empiric basis for building a conversion table for voice and for audiometry data. I hope to develop this possibility in a succeeding paper.

The necessity of taking actual measurements of sound decay in a test room rather than relying on a theoretic prediction is illustrated in figure 1. Here are shown the data for three test rooms. At the 15

foot distance, the sound decay is, of course, zero; then, for any shorter distance the decay is read on the ordinate directly in decibels.

The solid curve in figure 1 is the theoretic prediction for free field; the open circles for our sound-proof room; the closed circles for our voice test alley, a wooden, painted room 26 by 5 by 9¼ feet (7.9 by 1.5 by 2.8 meters) high, and the crosses for a classroom 30 by 21 by 10 feet (9.1 by 6.4 by 3 meters) high. Present in the sound-proof room the classroom and the alley was a white noise registering respectively 50, 52 and 58 decibels on a sound level meter at the position of the ear. This ambient noise was introduced merely to provide a uniform mask. It was not loud enough to induce serious harmonic distortion in the acoustic analyzer and should have no effect on the shape of curves in figure 1.

The difference between sound decay in the free field and in the sound-proof room is never greater than 4 decibels at any distance. What is even more surprising is the close correspondence between the sound-proof room and the acoustically untreated classroom, the latter being only 3 decibels worse at most and then only for the very near distances. Throughout the greater part of the distance the two rooms yield indistinguishable results.

The data for the narrow test alley diverge much more from the free field equation. Relatively little sound decay occurs through 14 feet (4.2 meters)—only about 12 decibels.

The practical significance of these facts can easily be shown: if, for example, the ambient noise in a test chamber raises the threshold by 20 decibels, the ears suffering loss up to 20 decibels cannot be detected from normal; now with a voice intensity constant at $d_o = 15$ feet, figure 1 shows that even in a free field a range of only 29.5 decibels would be available, and this is reduced to only about 24 decibels in an actual room. It is true, therefore, that in a room causing a threshold shift of 20 decibels with a whispered voice constant at $d_o = 15$ feet only the ears with losses of 20 to 44 decibels could be placed at all reliably on an acuity scale. Persons with losses less than 20 decibels would be reported perfectly normal, though some would have difficulty hearing in other situations, and those with losses greater than 44 decibels would be uniformly classed as completely deaf, though many of them would have usable residual hearing.

We know, however, from a vast amount of clinical data from this and other laboratories, that testers can and do consistently screen out ears which show loss of less than 20 decibels and can distribute on an acuity continuum the ears which have worse than a 44 decibel loss.

It is obvious that to do these things some variable other than distance is used. The tester must have changed the intensity of his voice in some systematic fashion.

In such cases, in which a tester uses both distance and voice intensity as variables, the foregoing calculations of test reliability do not apply. It must be emphasized that the present conclusions as to the reliability and usability of a free voice test apply only to a normal or near-normal population and not to a population of seriously deafened ears. For such testing, the error of prediction will be greater; in order to determine the amount of increase of error, further calculations would be necessary.

D. Voice Intensities Necessary for an Effective Free Voice Test.—From an article by E. P. Fowler Sr.,[15] it is possible to calculate the voice intensities required in order to use the voice as an instrument for testing seriously deafened ears. Free voice distance scores are given for 17 ears, together with complete audiograms. From the distance at which the voice could just be heard and from the audiometric hearing loss, the voice intensity used can be calculated in terms of how far a normal ear could hear that voice. The assumptions on which this calculation is based are all well documented in the present paper.

Figure 2 shows the variation in voice intensity used for the seriously deaf. Although only 1 to 4 ears are included at each distance, a fairly regular tendency is seen for the test voice to become louder as the ears tested are poorer in acuity. The tendency may be summarized by stating that the log of the voice intensity varies inversely with the voice score. In absolute terms, the intensity used to produce a 20 foot (6 meters) voice score could be heard by a normal ear at 43 feet (13 meters) (log 1.6335), and at the other extreme the intensity used to produce a 1 foot score could be heard at 2015 feet (614 meters) (log 3.3043).

It is evident that the accurate placement of the voice at intensities from a low whisper to a loud shout is beyond the ability of any but the most experienced tester and that one could not reasonably expect any reliability on such a test with routine testing personnel.

E. Relative Testing Times.—Another measure of the relative efficiency of the two tests is the amount of testing time per man. Obviously in clinical work the element of time is of secondary importance, but there are activities processing thousands of men in which a saving of even half a minute is of significance.

The longest testing time is taken by individual audiometry. A procedure in which each threshold is crossed at least twice before the final reading is taken being used, an average of 4.5 minutes is needed to perform a 6 octave audiogram on the 2 ears of a subject of average or better intelligence and near-normal acuity. This figure will predict

15. Fowler, E. P., Sr.: Hearing Standards for Acceptance, Disability Rating and Discharge in the Military Services and in Industry, Laryngoscope **51**:937-956, 1941.

rather well how many such audiograms can be completed in a specific period.

For the same type of subject, the whispered voice individual examination as we routinely administer it takes thirty-four seconds for the 2 ears. This mean time includes a range of ten to one hundred seconds for several dozen subjects.

Two modifications of audiometry are available to effect considerable saving of time, the sweep frequency technic and group testing. Either of these will prove more economical of time than the free voice examination under certain conditions. When all that is wanted is assurance that an ear can hear within 10 decibels of normal at all frequencies, it is sufficient to set the attenuation dial at 10 decibels and then to

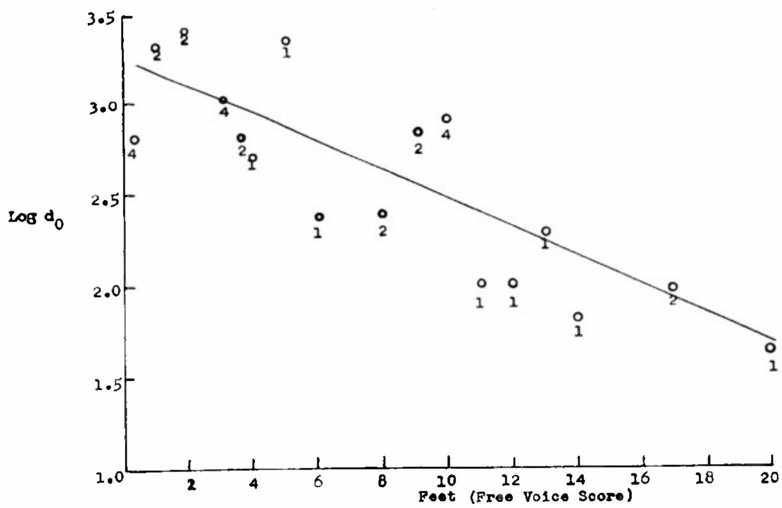

Fig. 2.—Relation of voice intensity to free voice score (calculated from data of E. P. Fowler Sr.). Note that numerals indicate the number of cases. Log d_0 indicates the distance at which the normal ear could hear a voice which produced a certain free voice score. For example, a free voice intensity which resulted in a score of 10 feet (average audiometer loss of 4 subjects of 38.2 decibels) could be heard a distance of 813 feet (log 2.9101) by a normal ear. An intensity resulting in a score of 20 feet (average loss of 6.6 decibels) could be heard a distancc of 43 feet (log 1.6335) by a normal ear.

sweep through the desired frequency range, interrupting the tone at appropriate points. By this means a satisfactory 6 octave screening audiogram can be performed on both ears in an average time of one minute and thirty-eight seconds.

If phones comparable in output are obtainable, a group of them can be coupled to the output of almost any commercial audiometer and complete audiometry performed on a number of subjects simultaneously.[10] We have found that a 6 octave audiogram for both ears can be performed on 12 subjects in twenty minutes on the average, or about one

minute and forty seconds per man, but there is no reason except lack of space or equipment why any number of phones should not be used and testing time cut to less than a minute per man.

In comparison of these audiometer testing times with those for the free voice, it must be remembered that the two types of test are not quite comparable, since all audiometer times given here are for 6 octaves, while the voice test provides pertinent information only for 3 octaves. If audiometer times were on a 3 octave basis, a time saving of perhaps a third would be effected.

III. GRANTED THAT AT A PARTICULAR ACTIVITY EITHER TEST COULD BE SET UP SATISFACTORILY, SHOULD ONE OR BOTH BE USED?

Since a free voice test and a pure tone test may be considered as interchangeable for routine screening when both are given under conditions approaching ideal, it remains to consider which should be recommended for a particular activity.

It is obvious that for a complete survey, as in clinical work, or for the fitting of hearing aids both speech and pure tones must be used, or else in some cases important information will be overlooked. It is also perfectly obvious that for certain job requirements only an extended frequency range, pure tone audiogram will suffice. But for the vast majority of acuity testing—screening for the armed services, Veteran's Administration, civil aviation or industry, where all that is required is a fair ability to hear human speech—only one type of test is necessary. The decision for any such activity need be based only on which test can be set up most efficiently and reliably.

Enough data have been presented in this paper to provide a basis for decision in some cases. For example, if group audiometry is possible for 36 subjects simultaneously, it will result in saving of time even over the most cursory free voice test and provide, in addition, valuable information on hearing at frequencies above and below the speech range. If audiometry must be individual, the free voice is quicker and under good conditions can, by careful supervision, be made almost as satisfactory as audiometry for the screening of a near-normal population. However, if the usual audiometric procedure is modified to a screening technic, whereby the operator merely assures himself that the hearing of a subject is better than a certain critical level, then individual audiometry is almost as rapid as a free voice test if even a little care is taken with the latter. Again, if testing must be accomplished on a population suffering generally from defective hearing, then the present paper makes it clear that there are factors at work which sharply affect the reliability of a free voice test and, unless rigorously controlled, are capable of reducing its meaning nearly to zero. On the other hand, while the precision of audiometry is reduced somewhat

with a deafened population, this reduction is of the order of a few decibels at most and is not to be compared with the error of prediction of a free voice test carried out close to a subject's ear. With such a population, audiometry in preference to the free voice should be mandatory. Other applications of the present data can be made to fit other needs.

SUMMARY

A reference paper on auditory acuity testing is presented, designed to help the beginning worker reach a decision as to which test or tests he should institute and to provide a body of data against which the experienced worker may check his own procedures.

Data are summarized from this and other laboratories bearing on the relative usability of the free voice and the pure tone audiometer as tests of auditory acuity. It was found the consensus of most studies that a careful voice test and pure tone audiometry through the speech range (512 to 2048 cycles per second) are intimately connected and measure almost the same auditory function. A comparison of six studies reveals the relationship between the two tests to be a straight line function. Statements cited to the effect that speech and audiometry are not related are definitely contraverted.

Recent studies on the reliability of these tests show that routine audiometry can well be expected to be accurate to within 5 decibels and that for near-normal subjects the reliability of the whispered voice need not be less.

A distance of 15 feet (4.5 meters) can be depended on to be a fairly accurate "normal" distance for the whispered voice, regardless of room acoustics and noise level.

A novel method of measuring sound decay as a function of distance is presented. By means of this method, the acoustic characteristics important for a voice test were compared with free field conditions for three representative test rooms. It was found that the usual narrow, reflective alley is almost useless for acuity testing. It was determined that for a population of near-normal acuity a whispered voice test in a fairly nonreflective room could serve but that for a seriously deafened population the additional variable of voice intensity would have to be added to that of distance.

Data from Dr. E. P. Fowler Sr. are analyzed to show that for a deafened population it is necessary to vary the voice from an average whisper to a loud shout. The log of voice intensity varies inversely with distance score. It is concluded that such control is beyond what can be expected from routine testers.

A summary is given of relative testing times for several modifications of the acuity test.

A statement is given of the relative applicability of voice and pure tone testing for a variety of specific situations.

Printed and Published in the United States of America